Information Modelling: An analysis of the uses and meanings of associations

Prof Matthew West, Shell Information Technology International Ltd, UK

1. Introduction

Historically many data models have taken a snapshot view of the world, which means that when change takes place, history is lost because it is overwritten. EPISTLE (the European Process Industries STEP Technical Liaison Executive) has been concerned with developing data models that are capable of managing information about process plants throughout the life of the plant. This means being able to hold information about the past, the present and the future.

The EPISTLE Core Model (ECM) has gone through a number of versions. In the early versions, history was supported by the use of associations. An association is a way of holding historical information about a relationship that can change over time. The relationship (in the entity-relationship sense) is replaced by an entity type that includes attributes that give the start date and end date for the validity of the relationship. This type of entity type is called an association. This approach is described in (1).

More recently, the way that change is handled in the EPISTLE Core Model has itself been changed. The approach now taken is to manage change through recognising different states of individual things that are valid for a period of time, together with timeless relationships between these. An outline of this approach is presented in (2). The resulting data model can be found in ECM V4.0 (3) and ISO/DIS15926-2 (4). The reasons for this change are beyond the scope of this paper. However, this paper does look at how the two approaches relate to each other. In particular we have found four basic patterns for how associations are translated into states and timeless relationships, and an approach to analysing associations is presented.

Note: the names of entity types used in this paper are not necessarily those to be found in ISO15926-2 or the EPISTLE Core Model.

2. From snapshot relationships to associations

In Figure 1 below, two ways of modelling the ownership of a physical object by an organisation are presented. The first models ownership as a relationship between an organisation and physical object. The problem comes when ownership changes. Either the change is not allowed, or the existing ownership attribute on the physical object is overwritten. This means that the previous ownership history is lost.

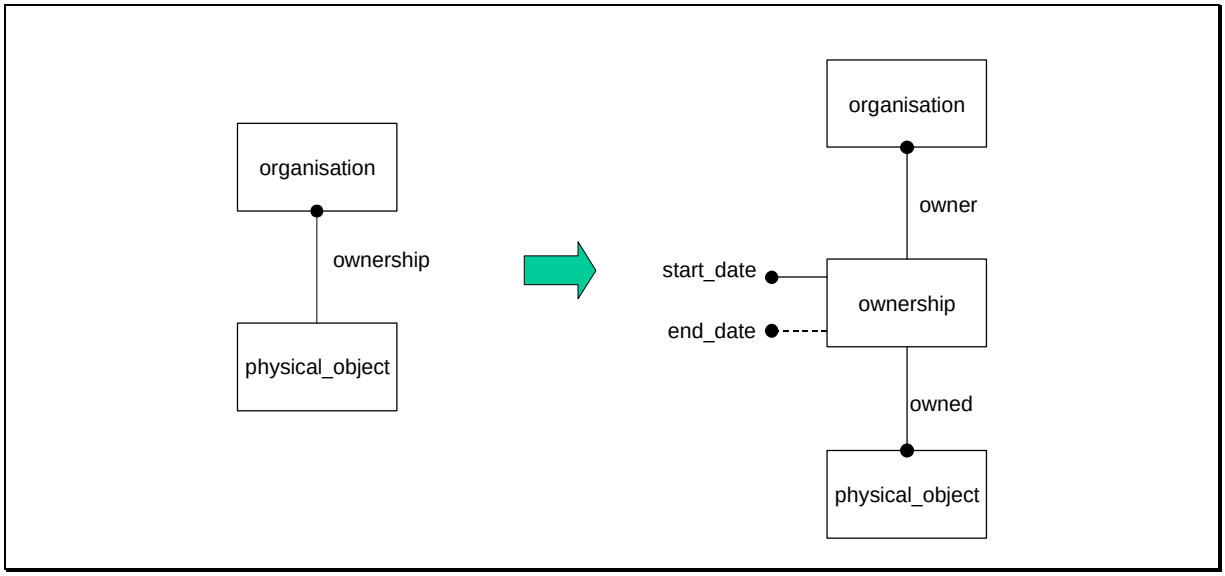

Figure 1: Moving from a snapshot model to one using associations.

The second models ownership as an association entity type. This has a start date and an end date as attributes, as well as references to what is owned, and the owner. Now when the ownership changes, the end date of the previous ownership relation gets an end date, and a new ownership record is created. As a result the ownership history can be held.

3. From associations to relationships between states and classes

Whilst using associations seems to solve the problem, further analysis suggests that some information is being hidden. In order to uncover this we need to consider individual things as spatio-temporal extents, and states as temporal parts of these spatio-temporal extents for which some relationship is true. To help in this we use space-time maps as illustrated in Figure 2 below.

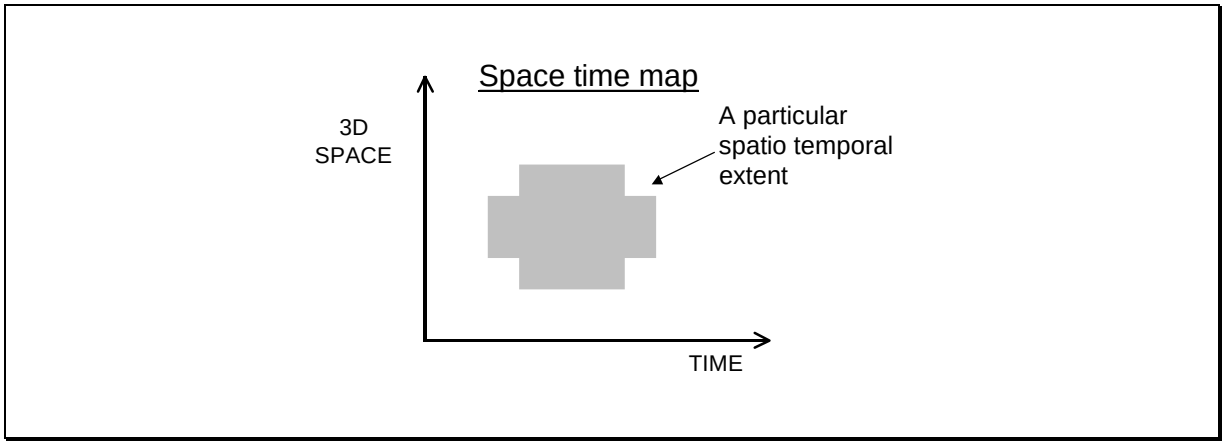

Figure 2: Space-time map.

Five different patterns have been found of what different sorts of association represent in spatio-temporal terms. These are illustrated in the following subsections.

3.1 Pattern 1: a relationship between a state of an individual and a class

Figure 3 shows a classification pattern. An example is given of where an individual_state, "Car1", is classified as being "Red" from 1/1/2001 to 4/3/2001.

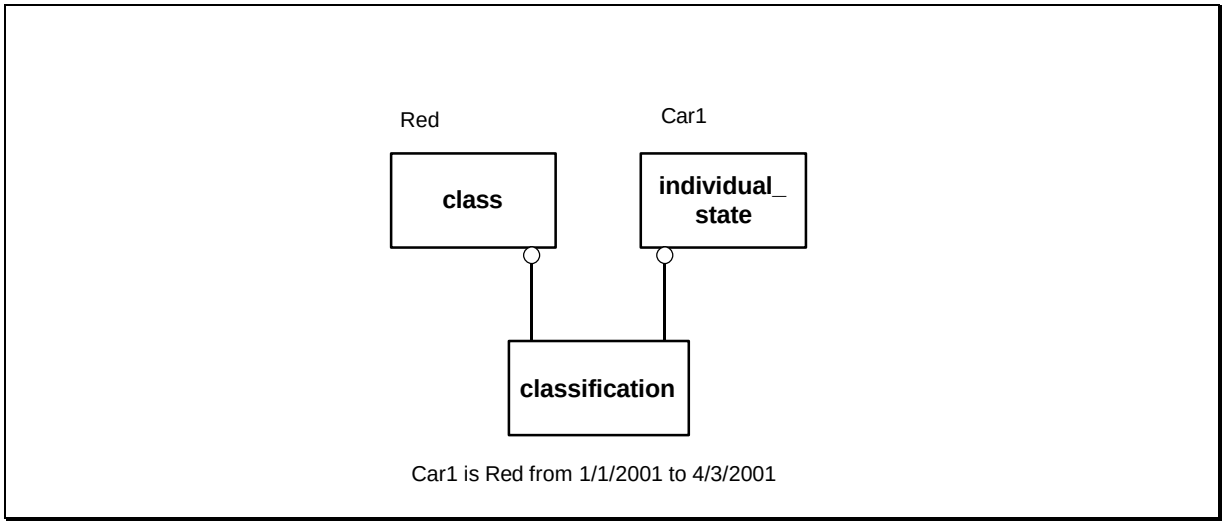

Figure 3: An example of a classification association.

If we examine what is happening here using a space-time map, Figure 4, we see that there is a state of Car1 that is classified as being red (the shaded area).

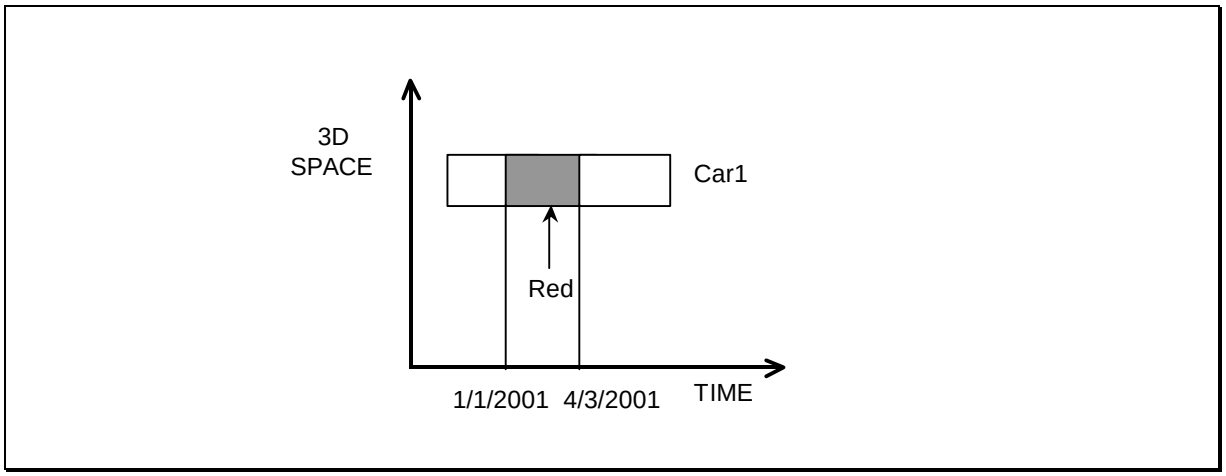

Figure 4: A space-time map for classification of an individual.

A data model that represents this space time diagram is shown in Figure 5.

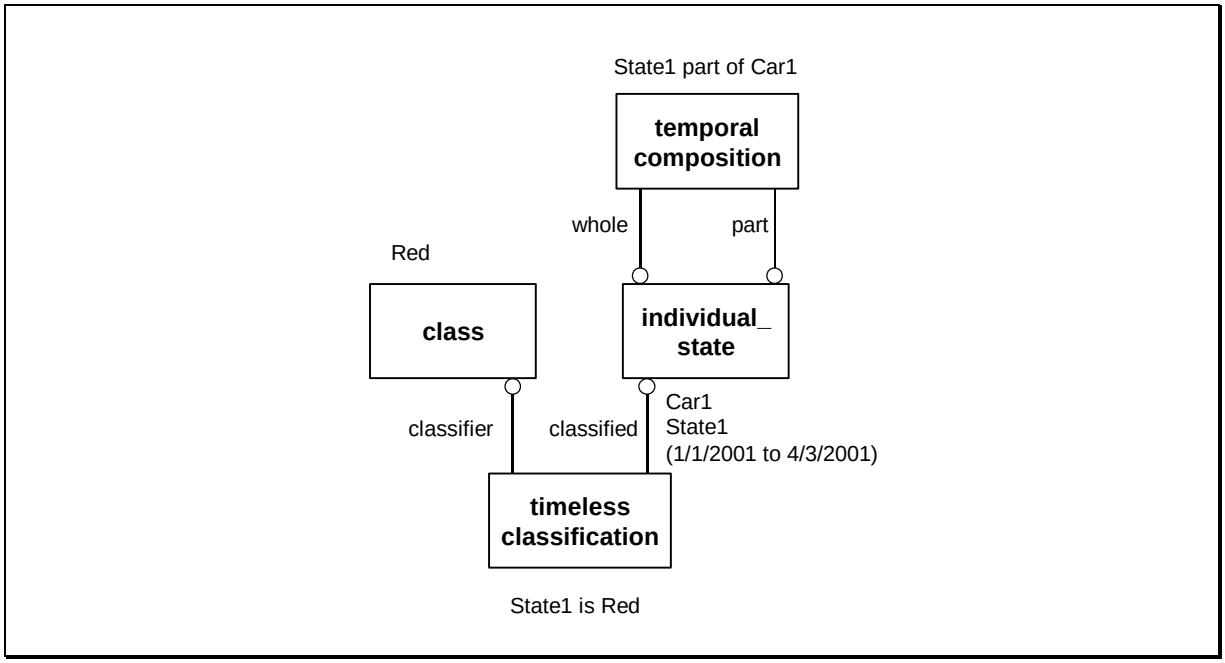

Figure 5: Classification using states.

Here the state of the car is modelled explicitly, rather than being part of the classification association, and it is shown as being a temporal part of the whole car, and as being classified as red. The classification relationship is now timeless, because the period of being red is explicit in the state that is classified.

3.2 Pattern 2: a relationship between two states of an individual

Figure 6 below illustrates the case where an association represents a relationship between two states of individual things. To illustrate the model an example is given of how a wheel, "Wheel1", is part of a car, "Car1" from 1/1/2001 to 5/4/2001.

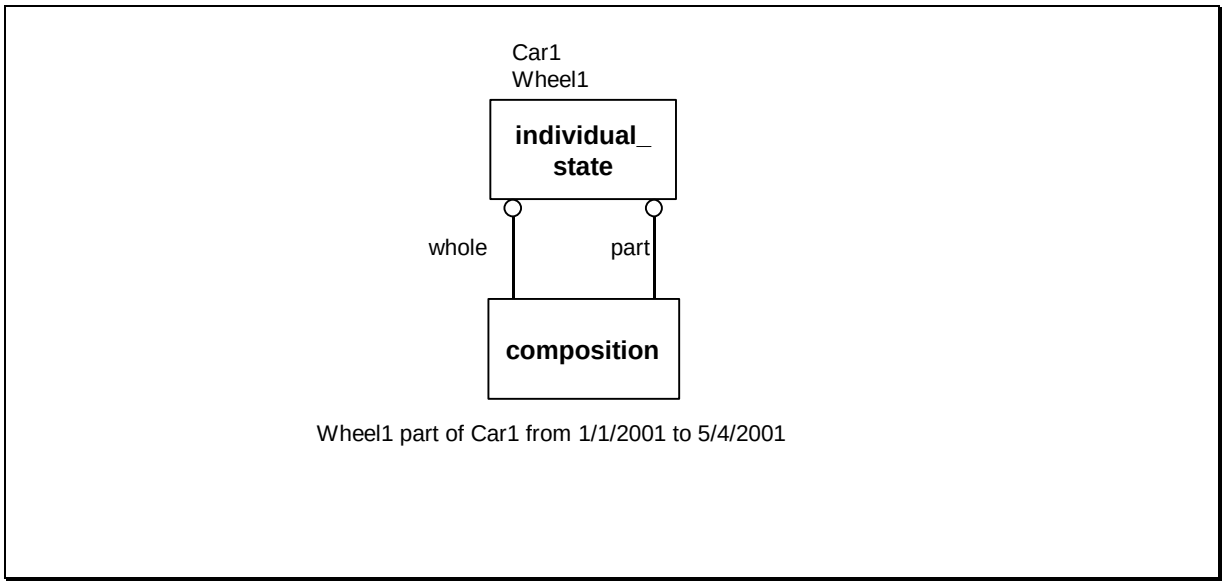

Figure 6: Association between two individuals.

Figure 7 below shows this example as a space time map, showing the different states of the car and wheel, as well as the whole life of the car and wheel.

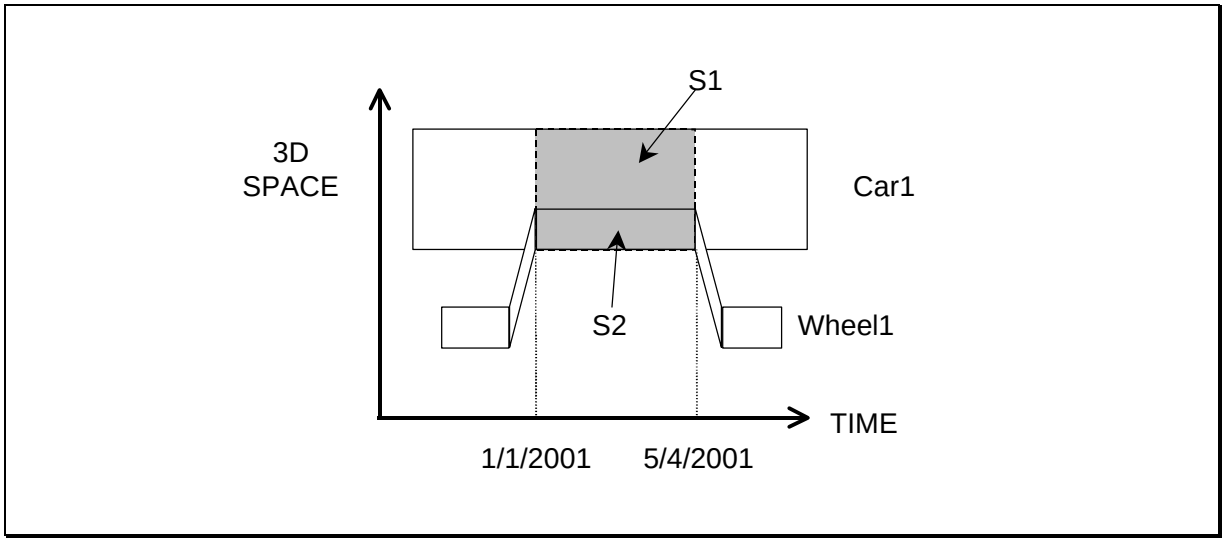

Figure 7: A space-time map for composition.

The diagram shows that in this case there is a state of Car1, S1, and a state of the Wheel1, S2, both with the same state and end date, and S2 is a part of S1. When this space-time map is modelled explicitly the result is found in Figure 8 below.

Here the individual_states S1 and S2 are modelled explicitly. S1 is shown as being a temporal part of Car1, S2 is shown as being a temporal part of Wheel1, and S2 is shown as being a part of S1.

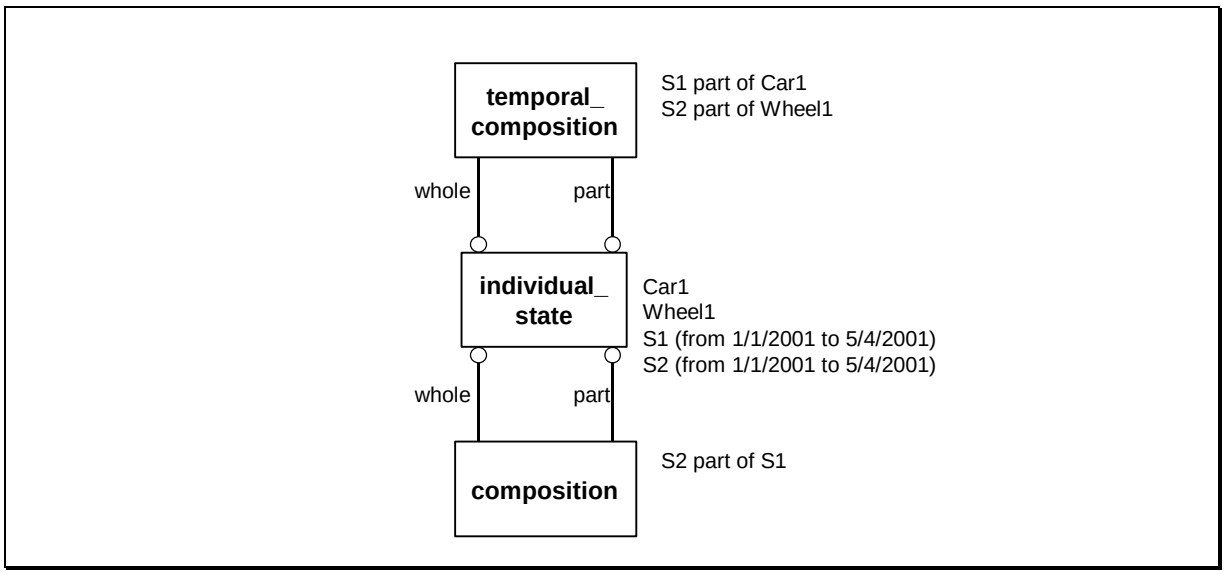

Figure 8: Composition using states.

3.3 Pattern 3: coincident individuals

Figure 9 below shows how a particular pump, Pump 1234, is installed to perform a particular duty, TAG P101, for a period of time between 3/1/2000 and 5/8/2001. At the end of this period, the pump is removed and replaced by another one performing the same duty.

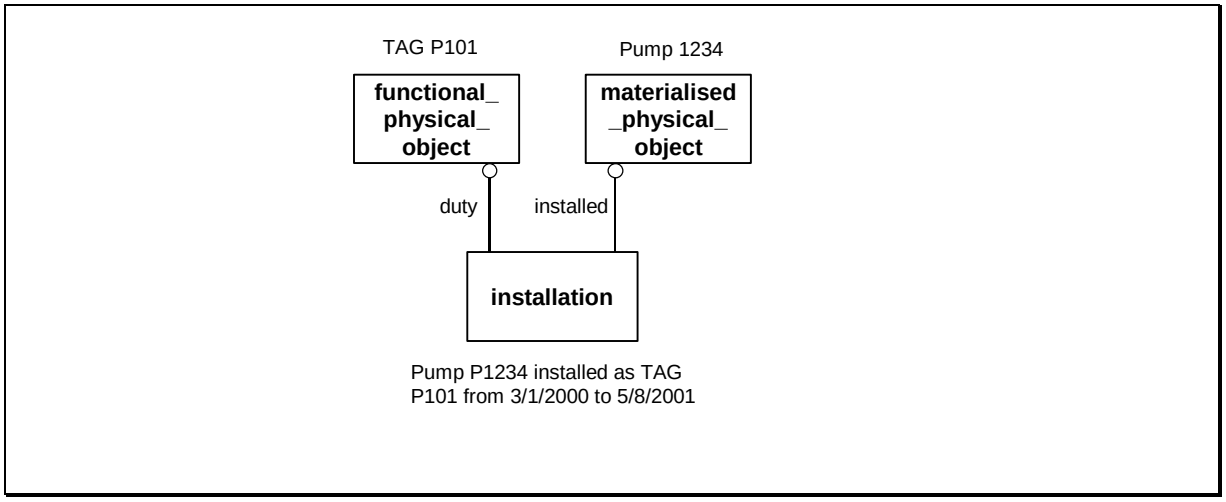

Figure 9: Coincident individuals.

Figure 10 below shows this example using a space-time map. Here it is possible to see that the duty represented by TAG P101, and Pump 1234 are coincident for the period of the installation, i.e. the state S1 of Pump 1234 that is installed as TAG P101 is in fact also a state of TAG P101. Indeed, the TAG P101 consists of those states of the pumps that are installed in this place.

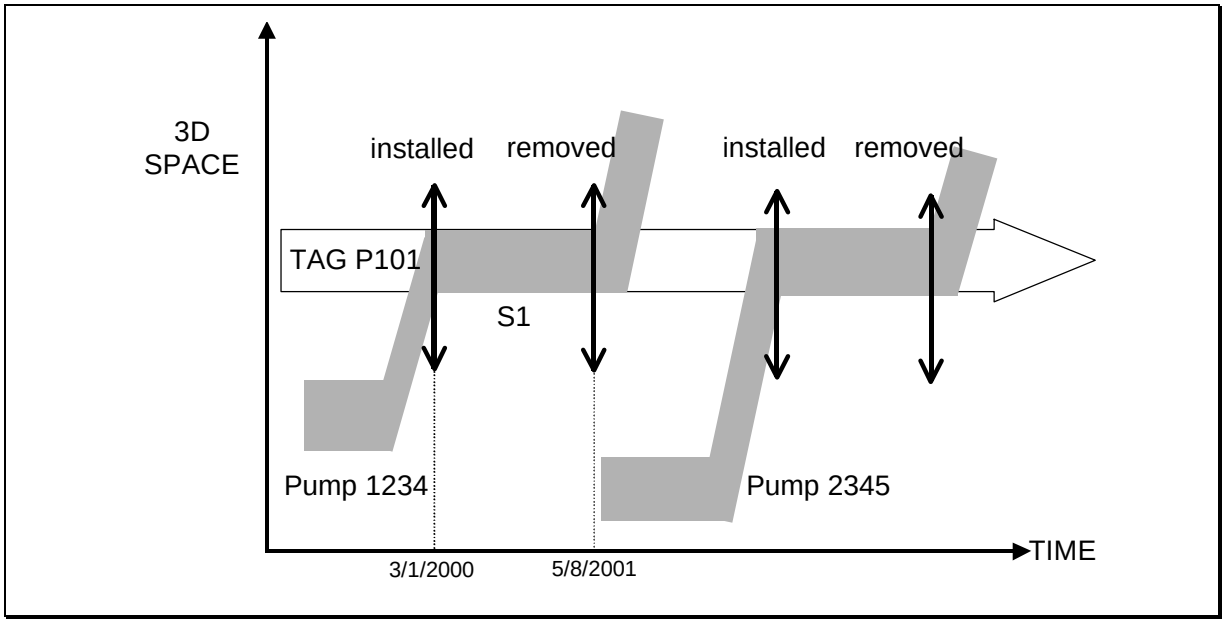

Figure 10: Space-time map for coincident individuals.

When this is modelled explicitly, rather than as an association, the model in Figure 11 results. Here S1 is shown as being a temporal part of both TAG P101 and Pump 1234.

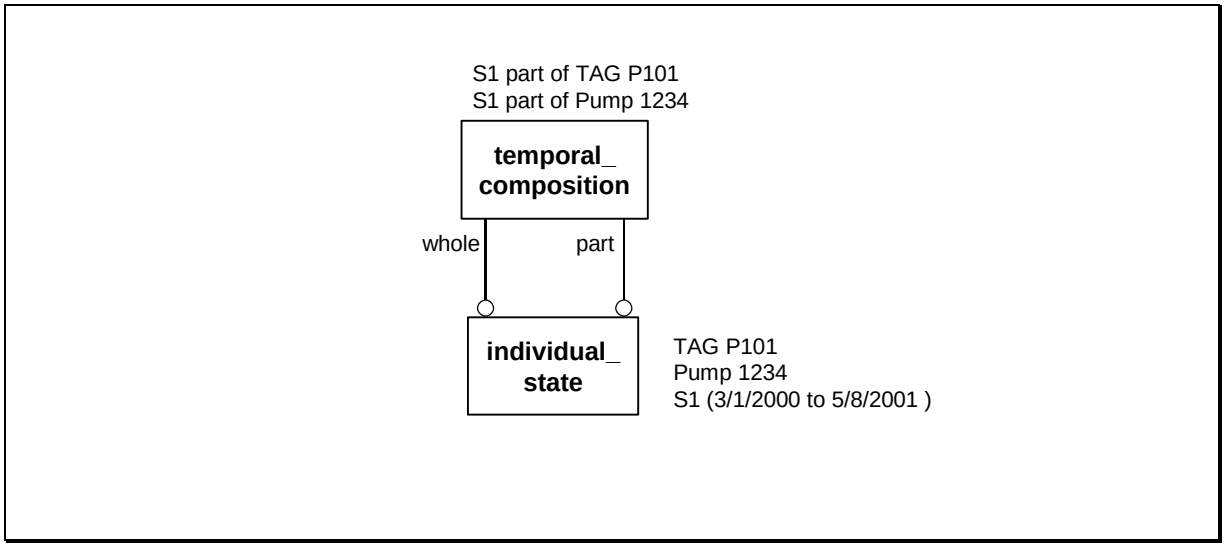

Figure 11: Coincident individuals using states.

3.4 Pattern 4: a relationship between two classes

Figure 12 below shows a specialisation association, and an example is given that says that a centrifugal pump is a specialisation of pump. The association automatically allows the specialisation to have a start and end date, but in fact classes are timeless, and there is no time when centrifugal pump is not a specialisation of pump.

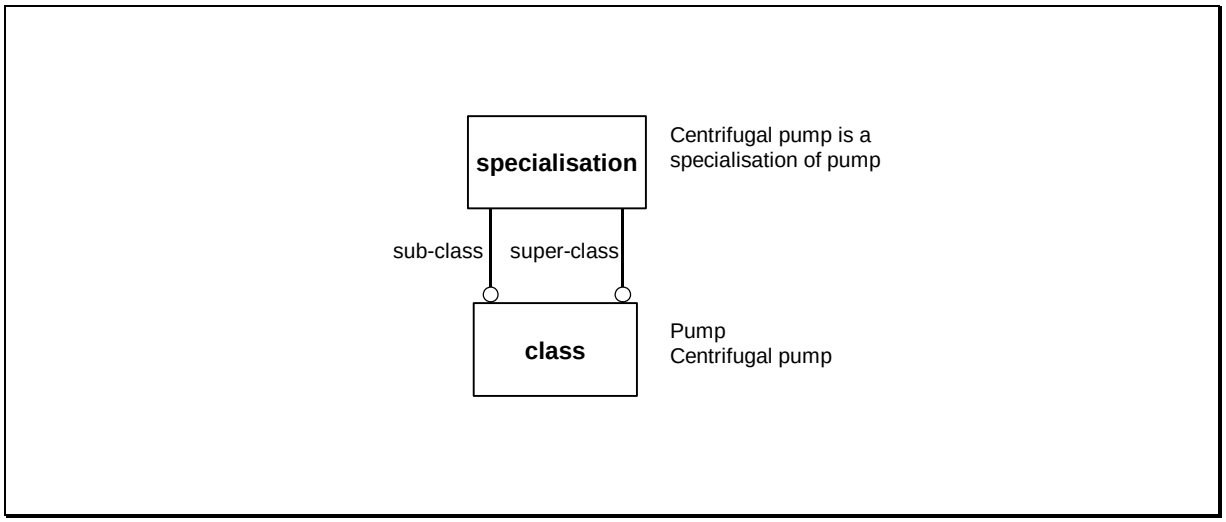

Figure 12: A relationship between two classes.

In this case there is no space-time map to draw, since all the objects sit outside of time. However, we can examine what the specialisation association means, and in this case we can determine that the specialisation association indicates that each member of the sub-class is also a member of the super-class. We therefore identify specialisation as a subtype of relationship as illustrated in Figure 13 below.

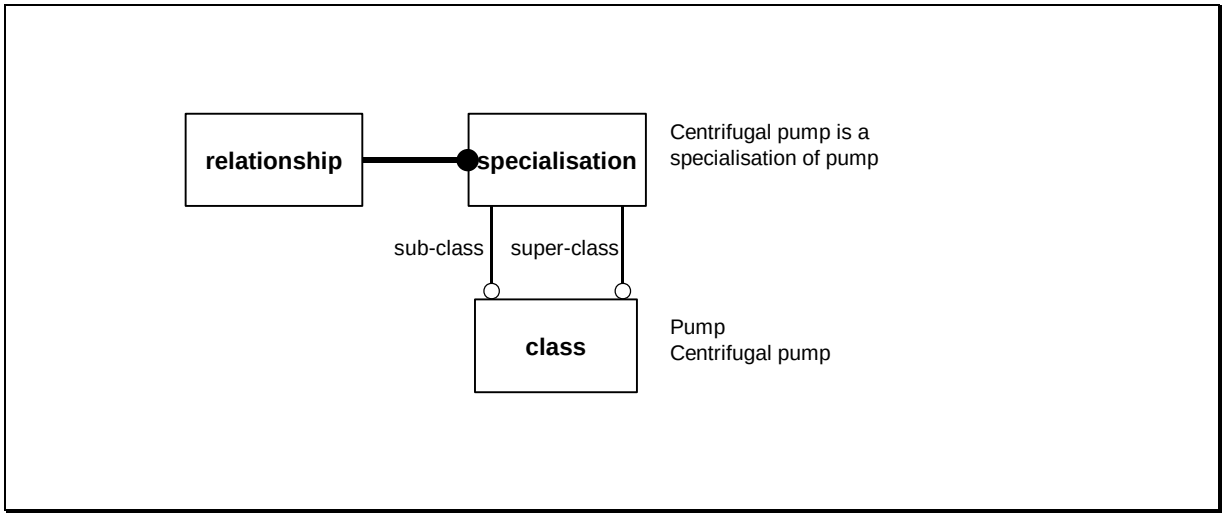

Figure 13: Analysis of a relationship between two classes.

3.5 Pattern 5: a class of relationship between two classes

Figure 14 shows an association between two classes that exhibits a different pattern. The association is composition_according_to_class, which is used to say that a member of the whole class, has a member of the part class as a part. The example given is that a centrifugal pump has an impeller as a part. Again, although an association can have a start and end date, none applies to this, as the statement is always intended to be true.

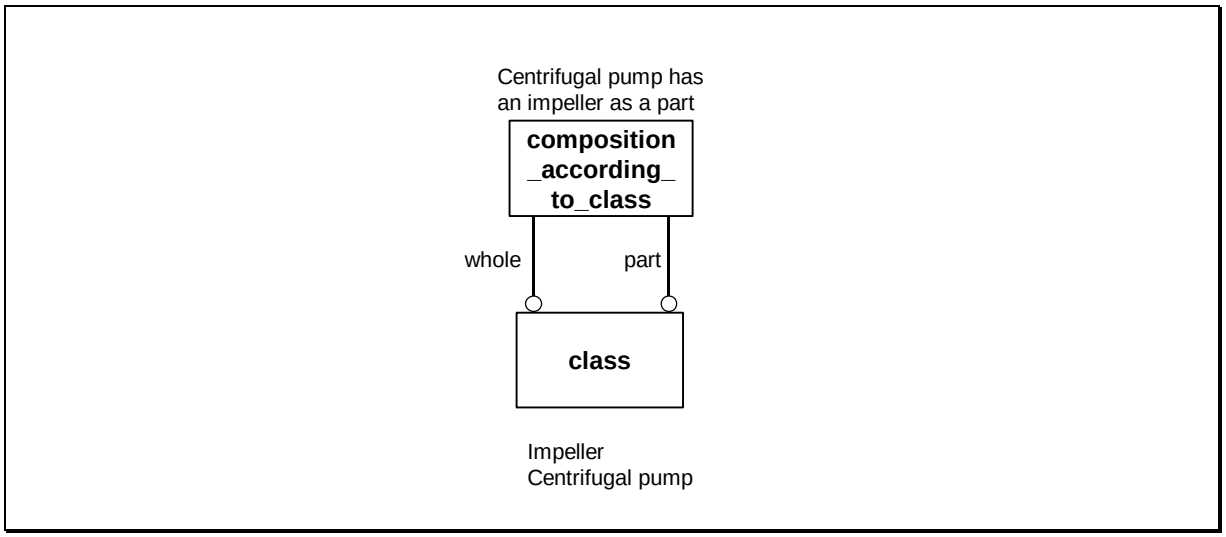

Figure 14: A class of relationship.

Once again there is no space time map, but analysis reveals a different pattern from the previous example. The association represents a rule again. In this case it is that each member of the whole class may have a member of the part class as a part. So each centrifugal pump may have an impeller as a part. Also in this case there will be particular relationships between particular impellers and particular centrifugal pumps that are instances of this rule. Hence, this association represents a class_of_relationship.

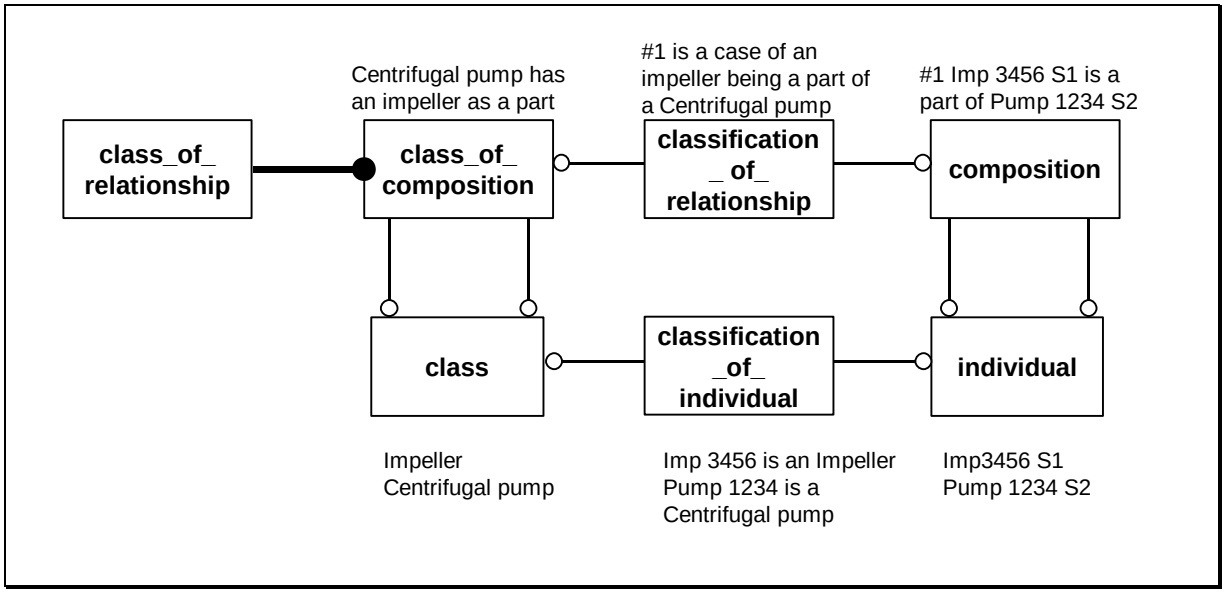

Figure 15: Analysis of a class of relationship.

This is illustrated in Figure 15 above. Here the composition relationship and the classification relationships are added that show how the rule applies to particular instances. An impeller state, Imp 3456 S1 is classified as an impeller; a pump state, Pump 1234 S2 is classified as a pump; a composition relation #1 shows that Imp 3456 S1 is a part of Pump 1234 S2; finally a classification of relationship shows that #1 is a case of an impeller being a part of a centrifugal pump.

4. Conclusions

Associations have been a powerful way to help in managing history, as opposed to just the current state of a domain. However, the analysis presented in this paper shows that considerable detail is hidden by the association construct. This paper shows that undertaking spatio-temporal analysis of associations allows a more precise model to be developed which makes this hidden detail more explicit. In particular five patterns of how associations can map into spatio-temporal terms have been identified and presented.

5. References

1. West, Matthew; Fowler, Julian; *Developing High Quality Data Models (Version 2.0)* EPISTLE (1996)

2. West, Matthew; *Some Notes on the Nature of Things* ISO TC184/SC4/WG10/N307 (2000)

3. *EPISTLE Core Model Version 4.0* EPISTLE (2001) http://www.epistle.ws/specifications/ecm.html

4. ISO/DIS 15926-2 - *Industrial automation systems and integration — Integration of life-cycle data for process plants including oil and gas production facilities — Part 2: Data model*. (2002)